

Investigations of the effects of random sampling patterns on the stability of generalized sampling

Robert Dahl Jacobsen* Jesper Møller* Morten Nielsen*
Morten Grud Rasmussen*

July 18, 2016

We investigate how the choice of spatial point process for generating random sampling patterns affects the numerical stability of non-uniform generalized sampling between Fourier bases and Daubechies scaling functions. Specifically, we consider binomial, Poisson and determinantal point processes and demonstrate that the more regular point patterns from the determinantal point process are superior.

1. Introduction

Generalized sampling [3, 2] is a technique for obtaining sparse approximations of signals in different bases and frames in a Hilbert space from their samples. The theory is abstract and does not restrict the type of bases or frames that can be considered, and this provides the freedom to adapt the setup to the structure of the signals at hand. In a typical application, we have samples of the signal given as inner products with respect to some fixed basis or frame imposed by the measuring process, but our knowledge of the signal structure dictates a desire to change to another more efficient representation system for the signal. For example, for images one may have access to Fourier samples of the image, but would like to change to a more efficient representation system such as wavelets.

An essential quantity ensuring numerical stability of the generalized sampling approach is the condition number of the change of basis (or frame) matrix between the two representation systems considered, see [2]. The condition number will, in general, depend on the sampling pattern and on the general structure of the representation system. If the sampling points are located on a regular grid, the numerical stability of the change of

*Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg East

basis matrix is well understood in the Fourier/wavelet setup [4]. However, the data acquisition process often forces one to consider non-regular sampling patterns which makes it essential to study numerical stability for non-regular sampling patterns, see e.g. [5].

In this paper we consider non-regular sampling patterns obtained randomly in the Fourier/wavelet setup and study how the condition number of the change of basis matrix can be controlled by using 'structured' sampling patterns. In general, any clustering of sampling points will result in large conditioning numbers and numerical instability. To avoid this problem, we generate sampling patterns from determinantal point processes where nearby points in the process repel each other in order to avoid clustering. We obtain samples from the binomial, Poisson and determinantal point processes (DPP) (see e.g. [14, 12]). The control of the corresponding condition numbers are obtained through certain geometric properties of the sampling pattern such as the density and bandwidth (as defined in Section 2.3). For the Poisson process, we derive a corresponding theoretical result (Theorem 5) relating the density and the intensity of the process. Estimating the probability that a random point patterns satisfies an appropriate density criterion is a natural extension of results in [2] for the parameters of non-random sampling schemes.

For other studies of numerical stability of random samples of Fourier systems, see [7, 17, 15].

2. Generalized sampling

In the following, we recap the framework of *generalized sampling*, introduced in a series of papers by Adcock, Hansen and collaborators, see [3, 1, 2]. Let \mathcal{H} be a separable Hilbert space and let $f \in \mathcal{H}$ be an element, we want to reconstruct from measurements. The measurements c_k are inner products $c_k = \langle s_k, f \rangle$, where $\{s_k\}$ is the set of sampling vectors. The reconstruction is an approximation of f of the form $\tilde{f} = \sum_j \beta_j r_j$, where r_j are called reconstruction vectors. In practice, the index sets for k and j should be finite, which means that the quality of the reconstruction $\tilde{f} \in \mathcal{W} := \text{span}_j \{r_j\}$ depends heavily on the choice of reconstruction vectors. In the following, we assume that the unknown function f is an element of $L^2(\mathbb{R}^d)$ supported in the compact set E , and we focus on frequency measurements (i.e. the c_k 's are certain values of the Fourier transform of f) and wavelet reconstructions (i.e. the r_j 's are Daubechies wavelets). Note that in this setup, the measurements can naturally be labelled by the frequencies at which the Fourier transform is evaluated.

2.1. Non-uniform generalized sampling

Let Ω denote the set of frequencies we measure, in the following referred to as a *sampling scheme*. If Ω is not a subset of a (sufficiently nice subset of a) lattice in \mathbb{R}^d , the classical theory is not applicable, and we need to use the so-called *non-uniform generalized sampling*, see [1, 2]. In the following, we recall some results from these papers.

Theorem 1 ([1])

Assume that we have a countable sampling scheme $\Omega \subset \mathbb{R}^d$, a finite dimensional recon-

struction space, i.e. a subspace $T \subset \mathcal{H} = L^2(\mathbb{R}^d)$, and a non-negative, bounded, selfadjoint operator $S: \mathcal{H} \rightarrow \mathcal{H}$ such that for all $f \in \mathcal{H}$, Sf is uniquely determined by $\{\hat{f}(\xi)\}_{\xi \in \Omega}$ and there exists a $C > 0$ such that for all $f \in T$, $\langle Sf, f \rangle \geq C\|f\|^2$. Then for every $f \in \mathcal{H}$ there exists a unique $\tilde{f} = F(f) \in T$ determined by $\forall g \in T: \langle Sf, g \rangle = \langle S\tilde{f}, g \rangle$ satisfying that $\forall f, h \in \mathcal{H}: \|f - F(f+h)\| \leq \sqrt{\frac{\|S\|}{C}}(\|f - Pf\| + \|h\|)$, where P denotes the orthogonal projection on T .

We remark that given the assumptions of the theorem, the reconstruction is *quasi-optimal*: $\|f - F(f)\| \leq \sqrt{\frac{\|S\|}{C}}\|f - Pf\|$ and *numerically stable*: $\|F(h)\| \leq \sqrt{\frac{\|S\|}{C}}\|h\|$.

To apply the theorem, we need to verify the assumptions. In [2], it is shown that if $T \subset L^2(E) \subset \mathcal{H}$ for a sufficiently nice, compactly supported set E , and if $\{x \mapsto \sqrt{\mu_\xi} e^{2\pi i \xi \cdot x} \mathbb{1}_E(x)\}_{\xi \in \Omega}$ with suitably chosen $\mu_\xi > 0$ is a so-called *weighted Fourier frame* for $L^2(E)$, then S can be chosen to be the corresponding *frame operator* (see Section 2.2), and they establish sufficient conditions on Ω which ensures that such a weighted frame exists. Since we want to go to finite subsets of Ω , choose labelling such that $\{\xi_n\}_{n=1}^\infty = \Omega$ and let $\Omega_N = \{\xi_n\}_{n=1}^N$ for positive integers N . With T and S as above, [2] establishes sufficient conditions for the corresponding *truncated frame operator* S_N to satisfy the conditions of Theorem 1. In this case, $F(f) = \tilde{f}$ is given by $\tilde{f} = \arg \min_{g \in T} \sum_{n=1}^N \mu_n |\hat{g}(\xi_n) - \hat{f}(\xi_n)|^2$.

2.2. Frames

Here, we recall some basic facts about frames and relate them to the setup of generalized sampling. A frame is a generalization of a basis of a Hilbert space, where the elements are allowed to be linearly dependent. More precisely, a frame of a separable Hilbert space \mathcal{H} is a collection of vectors $\{e_k\}_{k \in \mathbb{N}}$ such that there exists numbers $A, B > 0$ such that:

$$\forall f \in \mathcal{H}: A\|f\|^2 \leq \sum_{k \in \mathbb{N}} |\langle f, e_k \rangle|^2 \leq B\|f\|^2. \quad (1)$$

A and B are called the upper and lower frame bounds and can be chosen to be optimal. $A > 0$ implies that $\overline{\text{span}}_{k \in \mathbb{N}} \{e_k\} = \mathcal{H}$ while $B < \infty$ is needed for regularity and e.g. implies that the corresponding frame operator (see below) is a bounded operator and that (suitable) infinite linear combinations of the frame elements converges unconditionally.

The frame operator S of a frame $\{e_k\}_{k \in \mathbb{N}}$ is given by

$$Sf = \sum_{k \in \mathbb{N}} \langle f, e_k \rangle e_k.$$

In terms of the frame operator, the frame condition (1) can be written $A\|f\|^2 \leq \langle Sf, f \rangle \leq B\|f\|^2$, from which it follows that S is selfadjoint with positive upper and lower bounds and $\|S\| = B$ if B is the optimal upper frame bound.

The truncated frame operator S_N is just

$$Sf = \sum_{k=1}^N \langle f, e_k \rangle e_k.$$

In our context, the hope is that for a sampling scheme $\Omega = \{\xi_n\}_{n \in \mathbb{N}}$, $\{e_n\}_{n \in \mathbb{N}}$ given by $e_n(x) = \sqrt{\mu_\xi} e^{2\pi i \xi_n \cdot x} \mathbf{1}_E(x)$ is a frame for $L^2(E)$, where E is a sufficiently nice compact set (see below) and $\mu_\xi > 0$ (see below). In case it is, the frame operator S is just given by

$$(Sf)(x) = \sum_{n=1}^{\infty} \mu_{\xi_n} \hat{f}(\xi_n) e^{2\pi i \xi_n \cdot x},$$

and the frame condition is thus

$$A\|f\|^2 \leq \sum_{n=1}^{\infty} \mu_{\xi_n} |\hat{f}(\xi_n)|^2 \leq B\|f\|^2.$$

2.3. Generalized sampling with non-uniform sampling

Before we can state the sufficient condition for $\{e_n\}_{n \in \mathbb{N}}$ to be a frame, we need to introduce some notation.

Definition 2 ($|\cdot|_E$, E°)

Let E be compact, convex and symmetric. Then we define the norm $|\cdot|_E$ by

$$|x|_E = \inf\{a > 0 : x \in aE\}$$

and

$$E^\circ = \{y \in \mathbb{R}^d : x \cdot y \leq 1, \forall x \in E\}.$$

Note that E° is automatically also compact, convex and symmetric, so $|\cdot|_{E^\circ}$ is well-defined. We can now define the relative, inverse density of Ω in Y with respect to $|\cdot|_{E^\circ}$: $\delta_{E^\circ}(\Omega, Y) = \sup_{y \in Y} \inf_{\xi \in \Omega} |\xi - y|_{E^\circ}$. If the norm used is just the usual one, we omit the subscript “ E° ”.

Theorem 3 ([2])

If $\delta_{E^\circ}(\Omega, \mathbb{R}^d) < \frac{1}{4}$, then there exists weights $\mu_{\xi_n} > 0$, $\xi_n \in \Omega$ such that $\{e_n\}_{n \in \mathbb{N}}$ is a frame for $L^2(E)$. Moreover, the weights μ_{ξ_n} may be chosen as the measures of the Voronoi regions (of ξ_n) with respect to the $|\cdot|_{E^\circ}$ norm.

This result is obviously not applicable for finite sets Ω , as the inverse density of a finite set is infinite. But since we are only interested in vectors in the finite dimensional T , we can do with less than a frame for $L^2(E)$:

Theorem 4 ([2])

Assume that $T \subset L^2(E)$ is finite-dimensional, E is compact, convex and symmetric, and $\delta_{E^\circ}(\Omega, \mathbb{R}^d) < \frac{1}{4}$. Then $\{e_n\}_{n \in \mathbb{N}}$ with μ_{ξ_n} as in Theorem 3, is a frame.

Let A and B denote the frame bounds. Assume that K is a closed, simply connected set with 0 an inner point, such that $\Omega_N = \Omega \cap K$ is finite and that

$$R(\Omega_N, T) = \sup\left\{ \sum_{\xi \in \Omega \setminus \Omega_N} \mu_\xi |\hat{f}(\xi)|^2 : f \in T, \|f\| = 1 \right\} < A.$$

Then (T, Ω_N, S_N) satisfies the conditions of Theorem 1 with $C = A - R(\Omega_N, T)$ and $F(f) = \hat{f} = \arg \min_{g \in T} \sum_{n=1}^N \mu_{\xi_n} |\hat{f}(\xi_n) - \hat{g}(\xi_n)|^2$.

The set K in the above theorem is called the *bandwidth*. If the bandwidth is not sufficiently large, the reconstruction constant grows exponentially with the reconstruction space [1, Theorem 6.1].

3. Binomial, Poisson and determinantal point processes

In this section we assume Ω is a locally finite spatial point process in \mathbb{R}^d without multiple points and any accumulation point, i.e. we can view Ω as a closed random subset of \mathbb{R}^d so that with probability one, for any bounded set $B \subset \mathbb{R}^d$, the intersection $\Omega \cap B$ is finite. For measure theoretical details, see [14] and the references therein. We focus on three models: binomial, Poisson and determinantal point processes. We discuss only the definitions and properties of these models which become relevant for our purpose. In particular, in the case where Ω is a stationary Poisson process, we establish a lower bound on the probability that $\delta(\Omega, Y) < 1/4$.

We use the generic notation K for a Borel set $K \subset \mathbb{R}^d$ of finite Lebesgue measure $|K|$; e.g. K may be the observation window Y ; but the notation K will also be used for more general settings. Moreover, $N(K)$ denotes the random number of points in Ω falling in K .

3.1. Binomial and Poisson point processes

For a given Borel set $K \subset \mathbb{R}^d$ with $0 < |K| < \infty$, suppose Ω restricted to K consists of a fixed number $N(K) = n > 0$ of independent points ξ_1, \dots, ξ_n which are uniformly distributed on K (i.e. $P(\xi \in A) = |A|/|K|$ for Borel sets $A \subseteq K$). Then $\Omega \cap K = \{\xi_1, \dots, \xi_n\}$ is called a binomial point process.

The binomial point process can be extended to a stationary Poisson process as follows. Let $\rho > 0$. If for any Borel set $K \subset \mathbb{R}^d$ with $|K| < \infty$, we have that

- $N(K)$ is Poisson distributed with parameter $\lambda = \rho|K|$, that is $P(N(K) = n) = \exp(-\lambda)\lambda^n/n!$, $n = 0, 1, \dots$ (so if $|K| = 0$, then $N(K) = 0$),
- conditional on $N(K) = n > 0$, the n points in $\Omega \cap K$ form a binomial point process,

then Ω is a stationary Poisson point process on \mathbb{R}^d with intensity ρ . The process is often referred to as “complete spatial randomness” because of the lack of interaction or dependence properties.

The following theorem is verified in Appendix A, and it is of particular interest for us when $K = Y$ is the observation window.

Theorem 5

Suppose Ω is a stationary Poisson process with intensity $\rho > 0$ and consider any Borel

set $K \subset \mathbb{R}^d$ with $|K| < \infty$. Then

$$\begin{aligned} & \mathbb{P}(\delta(\Omega, K) < 1/4) \geq \\ & 1 - \rho^{d+1}|K| \frac{\Gamma\left(\frac{d^2+1}{2}\right)}{\Gamma\left(\frac{d^2}{2}\right)} \left\{ \frac{\Gamma\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d+1}{2}\right)} \right\}^d \frac{\Gamma\left(\frac{2}{2}\right) \cdots \Gamma\left(\frac{d}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \cdots \Gamma\left(\frac{d-1}{2}\right)} \frac{1}{d(\rho\omega_d)^d} \Gamma\left(d, \rho\omega_d 4^{-d}\right). \quad (2) \end{aligned}$$

3.2. Determinantal point processes

Determinantal point processes (DPPs) are models for repulsiveness (inhibition or regularity) between points in “space”, where space means \mathbb{R}^d in the present setting. DPPs are of interest because of their applications in mathematical physics, combinatorics, random-matrix theory, machine learning and spatial statistics, and because they provide rather flexible and tractable parametric models, see [12] and the references therein.

DPPs are defined in terms of the so-called joint intensities $\rho^{(n)} : \mathbb{R}^{dn} \mapsto [0, \infty)$, $n = 1, 2, \dots$: Recall that $\rho^{(n)}(x_1, \dots, x_n)$ (when it is assumed to exist) is only uniquely determined for Lebesgue almost all $(x_1, \dots, x_n) \in \mathbb{R}^d \times \dots \times \mathbb{R}^d$ (n times) is given by

$$\mathbb{E}\{N(A_1) \cdots N(A_n)\} = \int_{A_1} \cdots \int_{A_n} \rho^{(n)}(x_1, \dots, x_n) dx_1 \cdots dx_n$$

for pairwise disjoint Borel sets $A_1, \dots, A_n \subseteq \mathbb{R}^d$, where \mathbb{E} means expectation with respect to Ξ . Thus, for infinitesimally small A_1, \dots, A_n containing the points x_1, \dots, x_n , respectively, intuitively $\rho^{(n)}(x_1, \dots, x_n)|A_1| \cdots |A_n|$ is the probability that $\Omega \cap |A_i| \neq \emptyset$, $i = 1, \dots, n$.

Now, Ω is said to be a DPP with kernel $C : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{C}$ if we can take

$$\rho^{(n)}(x_1, \dots, x_n) = \det\{C(x_i, x_j)\}_{i,j=1,\dots,n}$$

for all $n = 1, 2, \dots$ and $(x_1, \dots, x_n) \in \mathbb{R}^d \times \dots \times \mathbb{R}^d$. The stationary Poisson process with intensity ρ is the very special case where $\rho^{(n)}(x_1, \dots, x_n) = \rho^n$, i.e. when $C(x, x) = \rho$ and $C(x, y) = 0$ whenever $x \neq y$.

In this paper we restrict attention to kernels defined by a continuous complex function $C_0 \in L^2(\mathbb{R}^d)$ so that $C(x, y) = C_0(x - y)$ is assumed to be positive semi-definite (then C is called a stationary covariance function). As discussed in [12] these are rather mild conditions, though the continuity assumption excludes the case of the Poisson process. For $\varphi \in L^1(\mathbb{R}^d)$, denote its inverse Fourier transform by $\mathcal{F}^{-1}(\varphi)$, i.e.

$$\mathcal{F}^{-1}(\varphi)(x) = \int \varphi(y) \exp(2\pi i x \cdot y) dy, \quad x \in \mathbb{R}^d.$$

Then the existence of the DPP is equivalent to the existence of a function $\varphi \in L^1(\mathbb{R}^d)$ such that $0 \leq \varphi \leq 1$ and $C_0 = \mathcal{F}^{-1}(\varphi)$, cf. Corollary 3.3 in [12]. Then φ is called the spectral density of the DPP, and the intensity $\rho^{(1)}(x) = \rho$ does not depend on x and is given by $\rho = C_0(0)$. Moreover, by comparison with the Poisson process, the repulsiveness of the process is reflected by that fact that $\rho^{(n)}(x_1, \dots, x_n) \leq \rho^n$. Therefore, we expect

that $P(\delta(\Omega, K) < 1/4)$ is larger than in the Poisson case. In fact this is in accordance with our experimental results discussed in Section 4. Moreover, an important observation in [12] is the trade-off between how large ρ can be and how strong the repulsiveness in the DPP can be.

Several examples of stationary covariance functions (Gaussian, Whittle-Matérn, generalized Cauchy and power exponential spectral model) and their spectral densities are discussed in [12]. A simulation algorithm for the corresponding DPPs is available in the **spatstat** package [6] — in particular with the contribution from [12].

For instance, a power exponential spectral model is specified by its spectral density given by

$$\varphi(x) = \rho (\alpha/\alpha_{\max})^d \exp(-\|\alpha x\|^\nu), \quad x \in \mathbb{R}^d,$$

where $\|\cdot\|$ denotes usual Euclidean distance, $\rho > 0$ is the intensity, $\nu > 0$ is a shape parameter and $\alpha \in (0, \alpha_{\max}]$ is a scale parameter, where

$$\alpha_{\max} = \sqrt{\pi} \left\{ \frac{\Gamma(d/\nu + 1)}{\Gamma(d/2 + 1)\rho} \right\}^{1/d}.$$

For fixed values of ρ and ν , the DPP becomes more and more repulsive as α increases to α_{\max} ; and for a fixed value of ρ and letting $\alpha = \alpha_{\max}$, as ν increases, the DPP ranges from the Poisson process (the limiting case of $\nu \rightarrow 0$) to the “most repulsive DPP” (the limiting case of $\nu \rightarrow \infty$), see [12] for the details.

4. Experiments

In this section we compare results obtained when using random sampling points generated from the binomial, Poisson and DPP models studied above. Recall that the binomial and Poisson point processes correspond to complete spatial randomness, whereas a DPP is repulsive, i.e. gives rise to more regular point patterns. This is seen in Figure 1 which shows a simulation from each process. Specifically, the DPP is given by a power exponential spectral model with shape parameter $\nu = 10$ and scale parameter $\alpha = \alpha_{\max}$; this is very close to the “most repulsive DPP”. Thus the Poisson point process and the DPP are specified by their intensity parameter ρ . The details are given in Section 4.1.

Simulations of the point processes are performed in R [16] with the **spatstat** package. Computations related to generalized sampling are performed in Julia [8] with the **GeneralizedSampling** package from [11] available at <https://github.com/robertdj/GeneralizedSampling.jl>. In this package the reconstructed scaling functions near the boundary are those of [9]. The code used in the experiments is available at <http://people.math.aau.dk/~robert/software>.

4.1. Setup

In the experiments, $d = 2$ and we have used a squared symmetric observation window $Y = [-64, 64]^2$, since the theory of generalized sampling dictates that the sampling area should be symmetric with a bandwidth that depends on the number of functions we wish

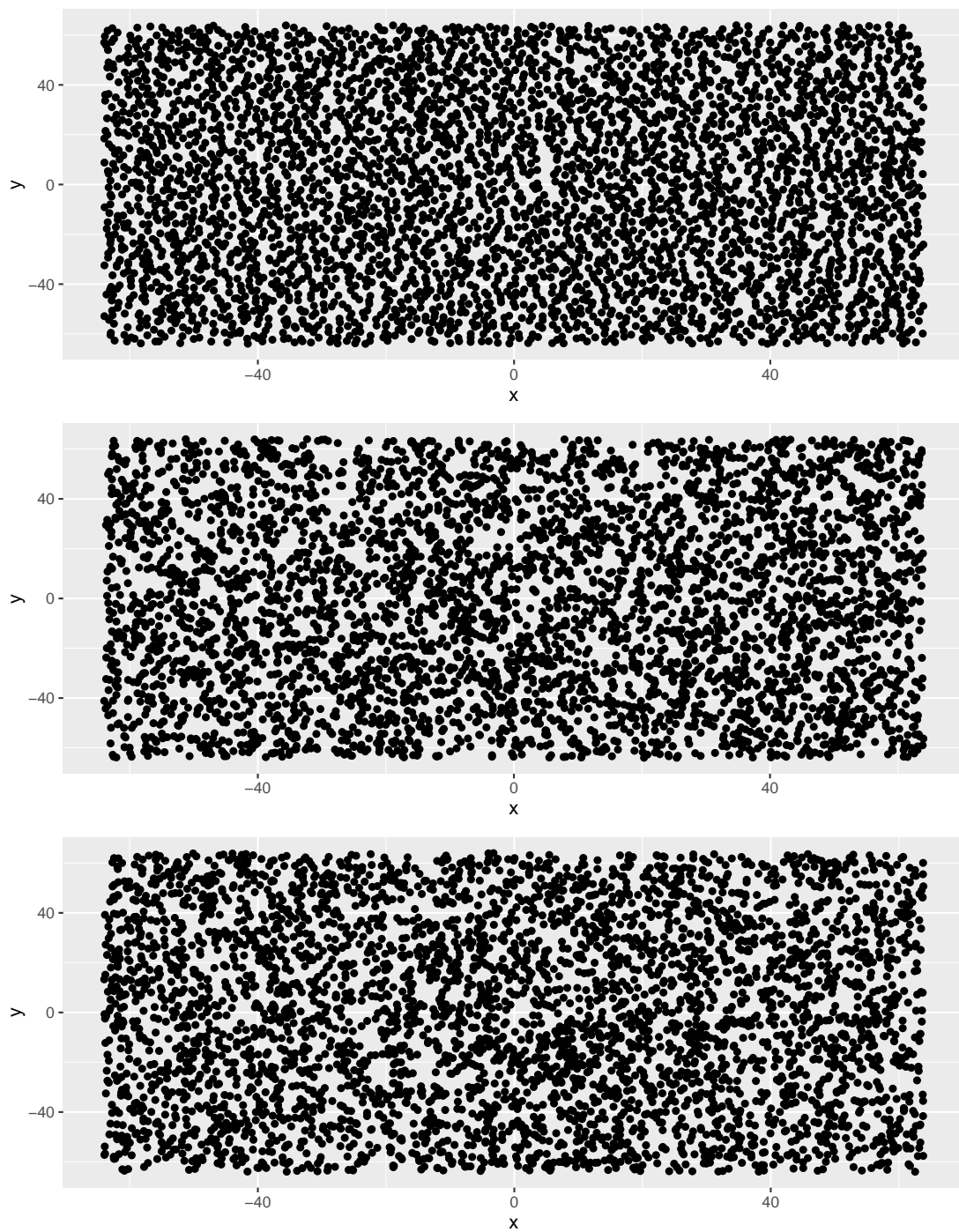


Figure 1: A realization of each of the point processes used in the experiments. From top to bottom the point patterns are from the DPP, Poisson and binomial point process.

to reconstruct. For each point process model we make 100 simulations, where the fixed number of points in the binomial point process is $n = 4096$, and the intensity of both the Poisson point process and the DPP is $\rho = 1/4$ so that the expected number of points is $E\{N(Y)\} = 4096$.

Depending on the nature of the sampling locations and the choice of reconstruction wavelet this allows us to recover *up to* $32^2 = 1024$ scaling functions, i.e. scaling functions at scale 5. However, this is overly optimistic in our experiments with non-uniform points that do not fulfill the density assumption of [2] so we have instead reconstructed at scale 4, i.e. with $16^2 = 256$ scaling functions. The support of the Daubechies scaling functions related to the wavelet with p vanishing moments at scale J is of length $2^{-J}(2p-1)$. Since our reconstruction space is $L^2([-1/2, 1/2])$ we must have less than or equal to 7 vanishing moments of the reconstructed wavelet to ensure that the support of the scaling functions are fully contained in $[-1/2, 1/2]$.

4.2. Results

We investigate three questions:

- Which point process is best for reconstruction (independently of the choice of Daubechies basis)?
- Which Daubechies basis is best for reconstruction (independently of the choice of point process)?
- How does the density of a point pattern influence the reconstruction?

In all cases we use the condition number of the change of basis matrix as a proxy for how much a certain factor influences the reconstruction. The condition number has a significant influence on how fast iterative solvers like conjugate gradient converges to the least squares solution we are interested in (see e.g. [10]).

The DPP consistently gives rise to the lowest condition numbers, see e.g. Figure 2 for the matrices used for reconstruction with the Haar scaling function. Strong evidence suggests that the mean condition number for the DPP is significantly lower than for both the binomial and Poisson point process (p-value $< 10^{-6}$).

For all three point processes the Haar scaling functions give rise to the lowest condition numbers, see e.g. Figure 3 for the matrices computed with points from DPP. The scales of the condition numbers are not comparable between the low and high order scaling functions, so boxplots for only the lower order scaling functions are seen in Figure 4. The mean condition number for the Haar scaling functions is significantly lower than for the remaining Daubechies scaling functions. This is consistent with the theoretical results for uniform generalized sampling in [4], i.e. where the sampling points are located on a grid.

It may be surprising that the increased regularity of the higher order Daubechies scaling functions does not appear to have a positive effect on the condition number. A possible explanation is that for the small number of reconstructed scaling functions we consider in the experiments the size of the support is more influential.

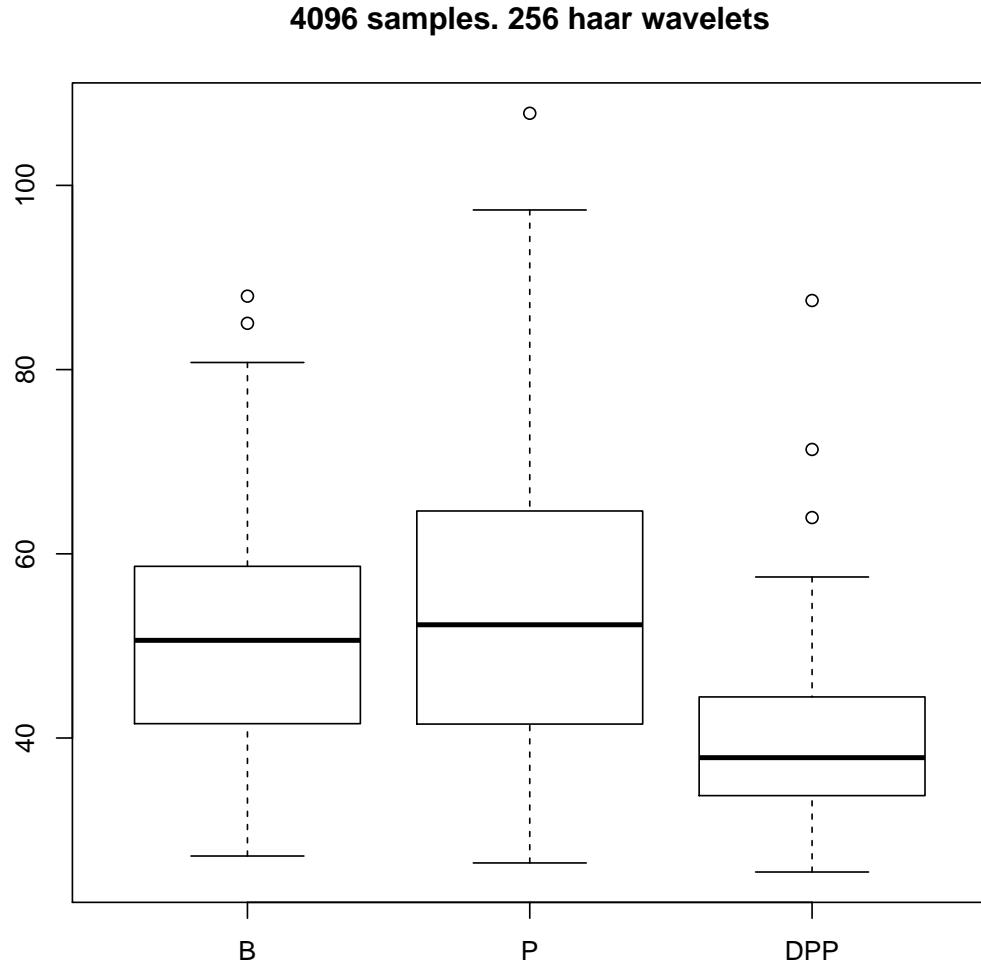


Figure 2: Condition numbers for change of basis matrices from 4096 samples to 16×16 Haar scaling functions.

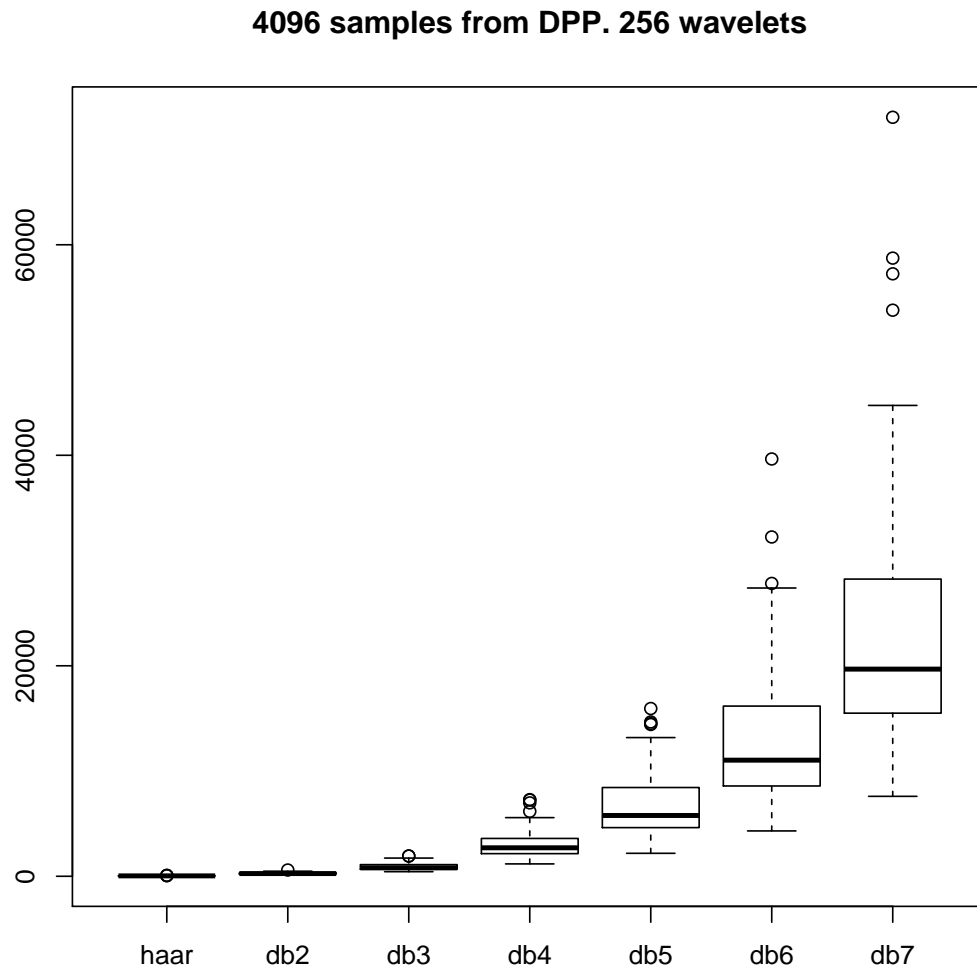


Figure 3: Condition numbers for change of basis matrices from 4096 samples from a DPP to 16×16 scaling functions.

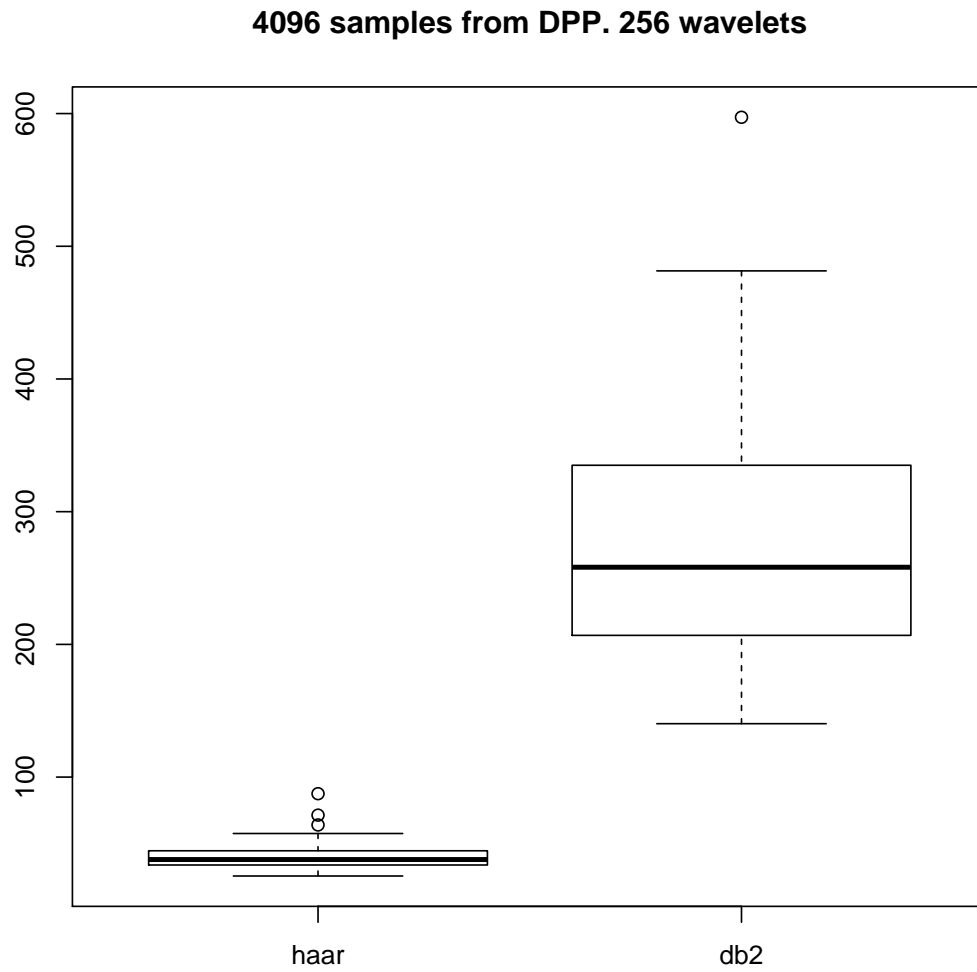


Figure 4: Condition numbers for change of basis matrices from 4096 samples from a DPP to 16×16 Haar and Daubechies 2 scaling functions.

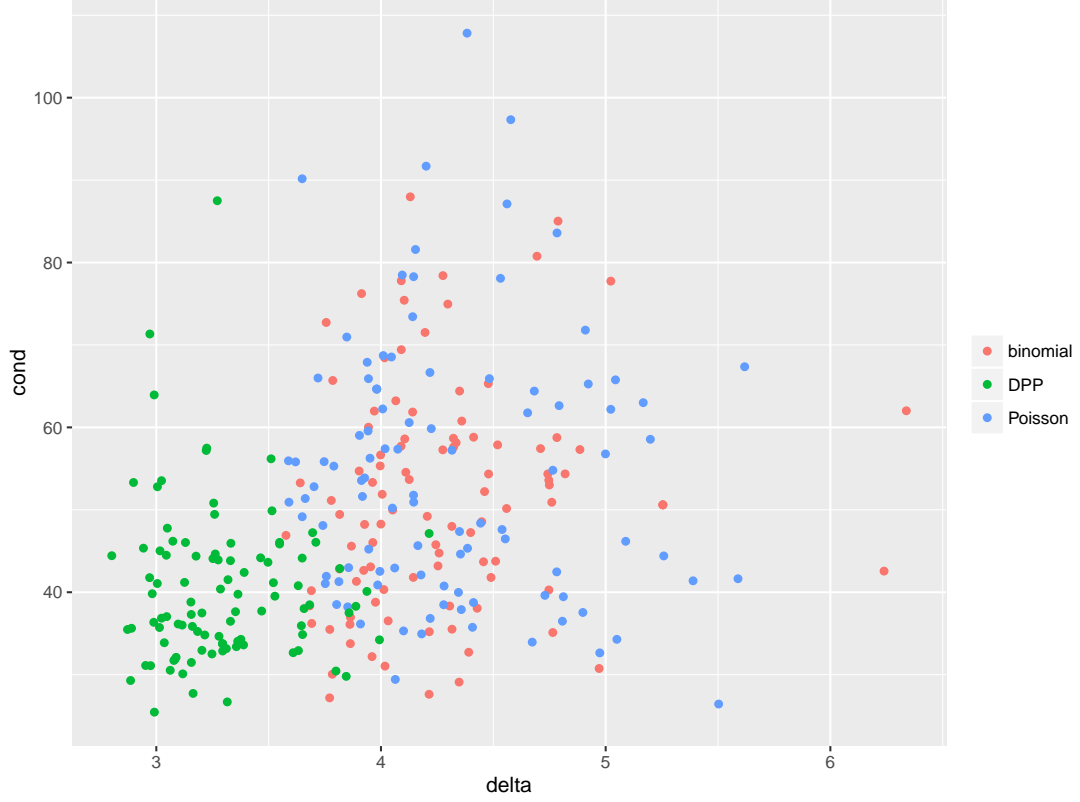


Figure 5: The density of a point pattern as defined in Section 2.3 versus the condition number of the change of basis matrix with sampling location from this point pattern to 256 Haar scaling functions.

As mentioned in Theorem 4 a small density of a point pattern is (part of) a sufficient condition for a small condition number of a change of basis matrix. In our simulations *none* of the realizations have small densities. Figure 5 shows scatterplots of densities of the point patterns versus the condition number of the change of basis matrix for reconstruction with the Haar scaling functions. For each of the processes *individually* there is no significant correlation between the density and the condition numbers (p-value ≈ 0.17 for the binomial point process, p-value ≈ 0.50 for the Poisson point process and p-value ≈ 0.72 for the DPP).

5. Conclusion

In this paper we have investigated how the choice of sampling according to a *random* point pattern affects the numerical stability of 2D non-uniform generalized sampling measured by the condition number of the associated change of basis matrix.

We have compared three kinds of random point patterns, generated by binomial,

Poisson or determinantal point processes. As reconstruction bases we have considered Daubechies scaling functions with polynomial reconstruction of degree 1 through 7 and the moment preserving boundary correction of [9].

Our results can be summarized as follows: 1) For all scaling functions sampling points from the determinantal point process yields the lowest mean condition number. 2) For all kinds of sampling points the Haar scaling function yields the lowest condition numbers. Point 2) is consistent with the findings of [4].

A. Proof of Theorem 5

Many statements in this appendix are only satisfied almost surely (a.s.), i.e. with probability one. They concern the stationary Poisson-Voronoi tessellation as studied in [13].

Let Ω be a stationary Poisson point process on \mathbb{R}^d with intensity $\rho > 0$. We refer to the points of Ω as nuclei and consider the Voronoi tessellation with cells generated by the nuclei, i.e. the Voronoi cell associated to a nucleus consists of all points in \mathbb{R}^d which are at least as close to that nucleus as to the other nuclei (with respect to usual distance in \mathbb{R}^d). With probability one each vertex in the Voronoi tessellation is given by intersection of $d+1$ Voronoi cells. If $\{\xi_0, \dots, \xi_d\}$ is the set of the corresponding nuclei, they define a.s. a unique d -dimensional closed ball $B(\xi_0, \dots, \xi_d)$ containing $\{\xi_0, \dots, \xi_d\}$ in its boundary. Its center $C(\xi_0, \dots, \xi_d)$ is then the Voronoi vertex. Conversely, for each set $\{\xi_0, \dots, \xi_d\}$ of $d+1$ pairwise distinct nuclei such that $B(\xi_0, \dots, \xi_d) \cap \Omega = \{\xi_0, \dots, \xi_d\}$, $C(\xi_0, \dots, \xi_d)$ is a.s. a Voronoi vertex. We denote $R(\xi_0, \dots, \xi_d)$ the radius of $B(\xi_0, \dots, \xi_d)$. Let $K \subset \mathbb{R}^d$ be a Borel set of finite Lebesgue measure $|K|$. Considering only those Voronoi vertices which are contained in K , the largest nuclei-vertex distance is a.s. given by

$$R(\Omega, K) = \max\{R(\xi_0, \dots, \xi_d) : \{\xi_0, \dots, \xi_d\} \subset \Omega \text{ is of cardinality } d+1, \\ B(\xi_0, \dots, \xi_d) \cap \Omega = \{\xi_0, \dots, \xi_d\}, C(\xi_0, \dots, \xi_d) \in K\}$$

noticing that the number of such vertices is a.s. finite. We want to estimate the probability $p = P(R(\Omega, K) < 1/4)$.

We have

$$1 - p = P\left(\exists \{\xi_0, \dots, \xi_d\} \subset \Omega \text{ of cardinality } d+1 : B(\xi_0, \dots, \xi_d) \cap \Omega = \{\xi_0, \dots, \xi_d\}, \right. \\ \left. C(\xi_0, \dots, \xi_d) \in K, R(\xi_0, \dots, \xi_d) \geq 1/4\right)$$

and this is at most the mean value

$$I = E \sum_{\{\xi_0, \dots, \xi_d\} \subset \Omega \text{ of cardinality } d+1} \mathbf{1} \left[B(\xi_0, \dots, \xi_d) \cap \Omega = \{\xi_0, \dots, \xi_d\}, \right. \\ \left. C(\xi_0, \dots, \xi_d) \in K, R(\xi_0, \dots, \xi_d) \geq 1/4 \right]$$

where $\mathbf{1}[\cdot]$ denotes the indicator function. Since Ω is a stationary Poisson point process we can evaluate I : Let

$$\omega_d = \frac{\pi^{d/2}}{\Gamma(1 + \frac{d}{2})}, \quad \sigma_d = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})},$$

be the volume respective surface area of the d -dimensional unit ball. By the extended Slivnyak-Mecke theorem (see e.g. [14]),

$$I = \rho^{d+1} \int \cdots \int \mathbf{1}[C(\xi_0, \dots, \xi_d) \in K, R(\xi_0, \dots, \xi_d) \geq 1/4] \\ \mathbb{P}(B(\xi_0, \dots, \xi_d) \cap \Omega = \emptyset) d\xi_0 \cdots d\xi_d$$

where

$$\mathbb{P}(B(\xi_0, \dots, \xi_d) \cap \Omega = \emptyset) = \exp(-\rho R(\xi_0, \dots, \xi_d)^d).$$

To evaluate this integral we make a shift of coordinates from (ξ_0, \dots, ξ_d) to (c, r, u_0, \dots, u_d) where $c = C(\xi_0, \dots, \xi_d)$, $r = R(\xi_0, \dots, \xi_d)$, and $x_i = c + ru_i$, $i = 0, \dots, d$. Let $\nabla = \nabla(u_0, \dots, u_d)$ be $d!$ times the Lebesgue measure of the simplex with vertices u_0, \dots, u_d . Then by the Blaschke-Petkantschin's formula (see [13] and the references therein),

$$d\xi_0 \cdots d\xi_d = \sigma_d^{d+1} \nabla r^{d^2-1} dc dr du_0 \cdots du_d$$

where dc is Lebesgue measure on \mathbb{R}^d , dr is Lebesgue measure on $(0, \infty)$, and du_i is the uniform distribution on the unit sphere in \mathbb{R}^d . Thus

$$I = \rho^{d+1} |K| \int_{1/4}^{\infty} r^{d^2-1} \exp(-\rho r^d) dr \int \cdots \int \nabla du_0 \cdots du_d.$$

Here

$$\int_{1/4}^{\infty} r^{d^2-1} \exp(-\rho \omega_d r^d) dr = \frac{1}{d(\rho \omega_d)^d} \int_{\rho \omega_d 4^{-d}}^{\infty} t^{d-1} e^{-t} dt = \frac{1}{d(\rho \omega_d)^d} \Gamma(d, \rho \omega_d 4^{-d})$$

where by integration by parts

$$\Gamma(d, s) = (d-1)\Gamma(d-1, s) + s^{d-1}e^{-s}, \quad s > 0.$$

Furthermore,

$$\int \cdots \int \nabla du_0 \cdots du_d = \frac{\Gamma(\frac{d^2+1}{2})}{\Gamma(\frac{d^2}{2})} \left\{ \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d+1}{2})} \right\}^d \frac{\Gamma(\frac{2}{2}) \cdots \Gamma(\frac{d}{2})}{\Gamma(\frac{1}{2}) \cdots \Gamma(\frac{d-1}{2})}$$

which reduces to 1 if $d = 1$ (see [13] and the references therein).

Consequently,

$$\mathbb{P}(R(\Omega, K) < 1/4) \geq \\ 1 - \rho^{d+1} |K| \frac{\Gamma(\frac{d^2+1}{2})}{\Gamma(\frac{d^2}{2})} \left\{ \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d+1}{2})} \right\}^d \frac{\Gamma(\frac{2}{2}) \cdots \Gamma(\frac{d}{2})}{\Gamma(\frac{1}{2}) \cdots \Gamma(\frac{d-1}{2})} \frac{1}{d(\rho \omega_d)^d} \Gamma(d, \rho \omega_d 4^{-d}).$$

Acknowledgments

Supported by the Danish Council for Independent Research | Natural Sciences, grant 12-124675, "Mathematical and Statistical Analysis of Spatial Data". Supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant (8721) from the Villum Foundation.

The authors would like to thank Ege Rubak for practical help with the `spatstat` package.

References

- [1] Ben Adcock, Milana Gataric, and Anders C. Hansen. On stable reconstructions from univariate nonuniform fourier measurements. *SIAM Journal on Imaging Sciences*, 2014.
- [2] Ben Adcock, Milana Gataric, and Anders C. Hansen. Weighted frames of exponentials and stable recovery of multidimensional functions from nonuniform fourier samples. *Applied Computational Harmonic Analysis*, 2015.
- [3] Ben Adcock and Anders C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *Journal of Fourier Analysis and Applications*, 18:685–716, 2012.
- [4] Ben Adcock, Anders C. Hansen, Gitta Kutyniok, and Jackie Ma. Linear stable sampling rate: Optimality of 2d wavelet reconstructions from fourier measurements. *SIAM Journal on Mathematical Analysis*, 47:1196–1233, 2014.
- [5] C. B. Ahn, J. H. Kim, and Z. H. Cho. High-speed spiral-scan echo planar nmr imaging-i. *IEEE Transactions on Medical Imaging*, 5(1):2–7, March 1986.
- [6] Adrian Baddeley, Ege Rubak, and Rolf Turner. *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, London, 2015.
- [7] Richard F. Bass and Karlheinz Gröchenig. Random sampling of multivariate trigonometric polynomials. *SIAM Journal on Mathematical Analysis*, 36:773–795, 2005.
- [8] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah. Julia: A fresh approach to numerical computing, 2014.
- [9] Albert Cohen, Ingrid Daubechies, and Pierre Vial. Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 1(1):54–81, December 1993.
- [10] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 4 edition, 2013.
- [11] Robert Dahl Jacobsen, Morten Nielsen, and Morten Grud Rasmussen. Generalized sampling in julia, 2016.

- [12] Frédéric Lavancier, Jesper Møller, and Ege Rubak. Statistical aspects of determinantal point processes. *Journal of Royal Statistical Society: Series B (Statistical Methodology)*, 77:853–877, 2015.
- [13] Jesper Møller. Random tessellations in \mathbb{R}^d . *Advances in Applied Probability*, 21:37–73, 1989.
- [14] Jesper Møller and Rasmus P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton, 2004.
- [15] Daniel Potts and Manfred Tasche. Numerical stability of nonequispaced fast Fourier transforms. *Journal of Computational and Applied Mathematics*, 222(2):655–674, 2008.
- [16] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.
- [17] Holger Rauhut. Random sampling of sparse trigonometric polynomials. *Applied and Computational Harmonic Analysis*, 22(1):16–42, 2007.